

# Problems with Blust's Phonology for Proto Austronesian

By Connor O'Donnell in 2025, updated in 2026

[www.parallax.net.nz](http://www.parallax.net.nz)

Blust has done a truly excellent job at reconstructing the Proto Austronesian language. However, to me, there are numerous extremely obvious discrepancies in his reconstructed consonant inventory, which is displayed below in a condensed format.

*/p/, */b/	*/t/, */d/	*/D/	*/k/, */g/, */j/	*/q/
*/m/	*/n/	*/ñ/	*/ŋ/	
	*/R/			
	*/r/			
	*/S/	*/s/		*/h/
	*/C/	*/c/, */z/		
*/w/		*/y/		
	*/l/	*/N/		

When converted into the standard international phonetic alphabet, it apparently becomes the following.

*/p/, */b/	*/t/, */d/	*/d/	*/k/, */g/, */gʲ/	*/q/
*/m/	*/n/	*/ɲ/	*/ŋ/	
	*/r/			
	*/r/			
	*/s/	*/ʃ/		*/h/
	*/ʈʂ/	*/tʃ/, */dʒ/		
*/w/		*/j/		
	*/l/	*/ʌ/		

To avoid confusing people who are already familiar with the subject, I will use Blust's transcription system instead of the international phonetic alphabet when referencing his reconstruction.

I will now go over each discrepancy.

## **\*/g/ and \*/j/**

These have mutually exclusive distributions, and the only difference in their realisations is that 1 is palatalised. Therefore, they're allophones. I would reconstruct them as \*/g/ which is allophonically realised as \*[j] intervocalically and in coda position.

## **\*/ñ/**

In Blust's dictionary, there are only 10 words that contain this phoneme, some of which are variants of the same word.

### **Group 1**

- ◆ \*/ñamñam/, \*/tañam/
- ◆ \*/ñamuR/
- ◆ \*/qañud/

### **Group 2**

- ◆ \*/ñawñaw/, \*/bañaw/, \*/Səñaw/, \*/Siñaw/
- ◆ \*/pəñu/
- ◆ \*/təñəb/

The words in group 1 appear to have the same outcomes in Formosan languages as they would have if they contained \*/N/ instead. For instance, Thao yields /z/ in its reflexes of both \*/ñamñam/ and \*/naNaq/, Atayal yields /l/ in its reflexes of both \*/tañam/ and \*/RiNuk/, and Saisiyat yields /l/ in its reflexes of both \*/qañud/ and \*/baNaS/.

The words in group 2 appear to have the same outcomes in Formosan languages as they would have if they contained \*/n/ instead. They appear to be reconstructed with \*/ñ/ in Proto Austronesian on the basis that there are a few Malayopolynesian reflexes that ended up with /ŋ/.


As we know, \*/N/ merged into \*/n/ in Proto Malayopolynesian. I suspect these 10 words were part of a small sporadic split during the Proto Malayopolynesian era, not before. All other instances of \*/ñ/ in Proto Malayopolynesian are in words that are not found in Formosan languages.

Therefore, based on the extremely low distribution, variability in Formosan reflexes, and the ability to produce the same outcomes with other existing phonemes, I would reconstruct /ñ/ as either \*/N/ or \*/n/ instead, depending on whether the word is in group 1 or group 2.

## **\*/N/**

In most Formosan languages, this became either /n/ or a type of lateral, in which case, it had a slight tendency to be palatalised. In Proto Malayopolynesian, it simply became \*/n/. To me, it seems that it could equally be reconstructed as either \*/ʎ/ or \*/ɲ/ in Proto Austronesian. I prefer to use \*/ɲ/ because it completes the row of nasals in the case that Blust's \*/ñ/ did not exist.

## **\*/s/ and \*/c/**

<b>PAN *susu<sub>1</sub> female breast, udder</b> 		
<b>Form.</b>		
Saisiyat	hœhœʔ	female breast
	ka-hœhœʔ-an	nipple of the breast
Hoanya	tutu	female breast
Sakizaya	cucuʔ	female breast
Thao	pa-tutu-an	nurse him! (e.g. crying child)
	tutu	breast
Bunun	susuʔ	female breast
	susu	breast
Bunun (Isbukun)	susuʔ	female breast
Saaroa	ʔu-susu	female breast
Amis	coco	breast of a female; udder; teat
	mi-coco	to suck or nurse; give the breast to
Puyuma	fufu	breasts
	ra-fufu-an	having big breasts
Paiwan	tutu	breasts

As seen in the image, in most Formosan languages, \*/s/ became either /s/, /t/, or /tʃ/. In Proto Malayopolynesian, it became /s/. The only evidence of palatalisation at this stage is in Puyuma as far as I can see. Based on those outcomes, I would reconstruct that it was originally \*/tʃ/.

In Blust's dictionary, there are only 4 words that contain \*/c/.

- ◆ \*/ləcɨk/
- ◆ \*/pəcək/
- ◆ \*/pəcəq/
- ◆ \*/qapucuk/

All of these appear to have the same outcomes in Formosan languages as they would have if they contained Blust's \*/s/ instead. They appear to be reconstructed with \*/c/ in Proto Austronesian on the basis that several Western Malayopolynesian reflexes ended up with /tʃ/. However, similarly to \*/ñ/, I suspect that these 4 words were part of a small sporadic split during the Proto **Western** Malayopolynesian era. All other instances of \*/c/ in Proto Western Malayopolynesian are in words that are not found in any Formosan language, and don't contain /tʃ/ in any other branch of Malayopolynesian as far as I can tell.

Therefore, I would reconstruct both \*/s/ and \*/c/ as \*/tʃ/. Separating them in Proto Austronesian is like trying to assert an underlying reason in Old English for the trap bath split to happen centuries later in a few dialects.

## ***\*/C/***

<b>PAN *Caliŋa ear</b> ↕		
<b>Form.</b>		
Bunun	taiŋa	ear
Saaroa	caliŋa	ear
Rukai (Budai)	caliŋa	ear
Rukai	caliŋa	ear
Rukai (Mantauran)	caliŋa	ear
Puyuma	Taŋila-ŋila-yan	the Jew's ear fungus, <i>Auricularia-Auricula Judae</i>
Paiwan	tsaliŋa	ear
	tsaliŋa-liŋa	<i>Coleus scutellarioides</i> ; <i>Iresine herbstii</i>
	tsaliŋa nua djaum	eye of a needle
Taokas	sanina	ear
Favorlang/Babuza	charina	ear

As seen in the image, in most Formosan languages, this became either */s/*, */t/*, or */tʃ/*. In Proto Malayopolynesian, it became *\*/t/*. The range of its outcomes are pretty much exactly the same as those of *\*/s/*, just in different places.

I would choose to reconstruct it as *\*/c/*, because it mirrors the outcomes of Blust's *\*/N/*. In Proto Malayopolynesian, *\*/N/* merged into *\*/n/*, and *\*/C/* merged into *\*/t/*. Those 2 defining shifts would be in harmony if *\*/N/* and *\*/C/* shared the same place of articulation and each had the same manner of articulation as the phonemes they became. It effectively turns it into 1 simple shift, which can be parsimoniously defined as "The alveolar palatal distinction was lost", instead of trying to explain how an affricate turned into a plosive and at the same time a palatal approximant turned into an alveolar nasal.

## **\*/z/ and \*/D/**

<b>PAN *zalan path, made by a human as opposed to an animal; way or means to do something</b> <span style="float: right;">↑</span>		
<b>Form.</b>		
Basai	tsatsan	path, road
Trobiawan	tsatsan	road
Kavalan	mani-razan	to take the same way
	nan-razan	same way of life
	razan	road, path, trail
	sa-razan	to pass through the bush
Saisiyat	raLan	road
Hoanya	dzalan	road
Sakizaya	daðan	path, road
Thao	saran	path, road
	mu-saran	go out onto the road; take out onto the road
Pazeh	darán	road, trail
Bunun	daan	road
	daan	road, path
Bunun (Isbukun)	daan	road
Kanakanabu	caáñə	road, path
Tsou	crənə	road
Amis	lalan	path, road, street, highway; direction figuratively; means to do something
Siraya	darəŋ	road, path
Puyuma	daʃan	road
	mi-a-daʃan	spirits of the road
	p<in>u-daʃan-an	place where a road has been made
Paiwan	djalan	road, trail, path
	ta-djalan a tsautsau	same kind of people
	pu-djalan	to have a road; make a road
	ta-djalan	same, same kind of ("one road")
Favorlang/Babuza	tarran	road, path

As seen in the image, in most Formosan languages, \*/z/ became either /d/, /tʃ/, or /r/. In Proto Malayopolynesian, it became \*/z/. The only evidence of palatalisation as far as I can see is in

Paiwan \*/t̚/. Based on those outcomes, I would reconstruct that it was originally realised as \*[dz] with a slight tendency to go in the postalveolar direction.

In Blust’s dictionary, there are only 3 words that contain \*/D/.

- ◆ \*/ikuD/
- ◆ \*/NataD/
- ◆ \*/SatəD/

Even in Proto Malayopolynesian, there are only 8. Additionally, \*/z/ and \*/D/ have mutually exclusive distributions. Therefore, they’re allophones, and the small distribution can be explained by the fact that \*/D/ is only 1 environment of a more widespread phoneme.

**PAN \*NataD outside** ↑

Form.		
Saisiyat	latar	outside
Rukai	látaɖə	outside
Rukai (Budai)	látaɖə	outside
Rukai (Tanan)	latáɖə	outside
Rukai	latáɖə	outside
Proto-Rukai	*lataɖə	outside

**PAN \*SateD accompany, escort; send, return** ↑

Form.		
Puyuma	ʔatəɖ	send, deliver
Paiwan	satjez	return something; object which is being returned

As seen in the images, in Formosan languages, \*/D/ became either /d/, /z/, or /r/. In most Western Malayopolynesian languages, it became either /d/ or /r/. From those outcomes, I find it hard to say if this phoneme was originally pronounced any differently in the \*/D/ environment, compared to the \*/z/ environment. So I would simply reconstruct them both as \*/d̥z/.

## My Resulting Reconstructed Consonant Inventory

*/p/, */b/	*/t/, */d/	*/c/	*/k/, */g/	*/q/
*/m/	*/n/	*/ɲ/	*/ŋ/	
	*/r/			
	*/r/			
	*/s/			*/h/
	*/t͡s/, */d͡z/			
*/w/	*/l/	*/j/		

### Allophones

- ◆ \*/g/ \*[j]      **Environment:** /V\_/

I believe that this is a much more accurate consonant inventory that also happens to be more consistent with the trends of natural human languages.

## Phonology Conversion Matrix

Blust's	Mine	My Suggested Orthography
*/g/	*/g/	⟨g⟩
*/j/		
*/n/	*/n/	⟨n⟩
*/ñ/		
*/N/	*/ɲ/	⟨ñ⟩
*/S/	*/s/	⟨s⟩
*/s/	*/t͡s/	⟨c⟩
*/c/		
*/C/	*/c/	⟨ɥ⟩
*/z/	*/d͡z/	⟨z⟩
*/D/		
*/R/	*/r/	⟨r⟩
*/r/	*/r/	⟨ř⟩
*/y/	*/j/	⟨y⟩
*/ŋ/	*/ŋ/	⟨ḡ⟩

## Resources

- ◆ **Blust's Dictionary:** [acd.clld.org](http://acd.clld.org)
- ◆ **My Dictionary:** [www.parallax.net.nz/proto\\_polynesian\\_cognate\\_dictionary](http://www.parallax.net.nz/proto_polynesian_cognate_dictionary)